Volume Title
*ASP Conference Series, Vol. **Volume Number***
Author

Calibrating the PAU Survey's 46 Filters

Anne Bauer[1], Francisco Castander[1], Enrique Gaztañaga[1], Santiago Serrano[1], Nacho Sevilla[2], Nadia Tonello[3] and the PAU Team

[1]*Institut de Ciències de l'Espai, CSIC/IEEC, F. de Ciències, Torre C5 par-2, Barcelona 08193, Spain*

[2]*CIEMAT, Avenida Complutense 22, E-28040, Madrid, Spain*

[3]*Port d'Informació Científica, Campus UAB, Edifici D, E-08193 Bellaterra, Spain*

Abstract. The Physics of the Accelerating Universe (PAU) Survey, being carried out by several Spanish institutions, will image an area of 100-200 square degrees in 6 broad and 40 narrow band optical filters. The team is building a camera (PAUCam) with 18 CCDs, which will be installed in the 4 meter William Herschel Telescope at La Palma in 2013. The narrow band filters will each cover 100Å, with the set spanning 4500-8500Å. The broad band set will consist of standard *ugriZy* filters. The narrow band filters will provide low-resolution (R$\sim$50) photometric "spectra" for all objects observed in the survey, which will reach a depth of $\sim$24 mag in the broad bands and $\sim$22.5 mag (AB) in the narrow bands. Such precision will allow for galaxy photometric redshift errors of 0.0035(1+z), which will facilitate the measurement of cosmological parameters with precision comparable to much larger spectroscopic and photometric surveys. Accurate photometric calibration of the PAU data is vital to the survey's science goals, and is not straightforward due to the large and unusual filter set. We outline the data management pipelines being developed for the survey, both for nightly data reduction and coaddition of multiple epochs, with emphasis on the photometric calibration strategies. We also describe the tools we are developing to test the quality of the reduction and calibration.

1. The PAU Survey

The Physics of the Accelerating Universe (PAU) Survey[1] is a narrow-band survey set to begin in 2013. The main science goals of the project are to study large scale structure up to redshift $\sim$1.4 by making precision measurements of the clustering of galaxies. The survey will include 6 broad band filters (*ugriZy*) and 40 100Å-wide narrow band filters from 4500Å to 8500Å (see Figure 1), which will enable the measurement of galaxies' redshifts with a precision of 0.0035(1+z) (an order of magnitude better than what is possible with broad band photometric data alone). This photometric redshift precision corresponds to a line-of-sight distance of 12.5 Mpc/h at a redshift of 0.5, allowing for the study of galaxy clustering in three dimensions at the length scales of linear structure growth. The main cosmology probes in the PAU Survey will be the study of redshift

[1]http://www.pausurvey.org

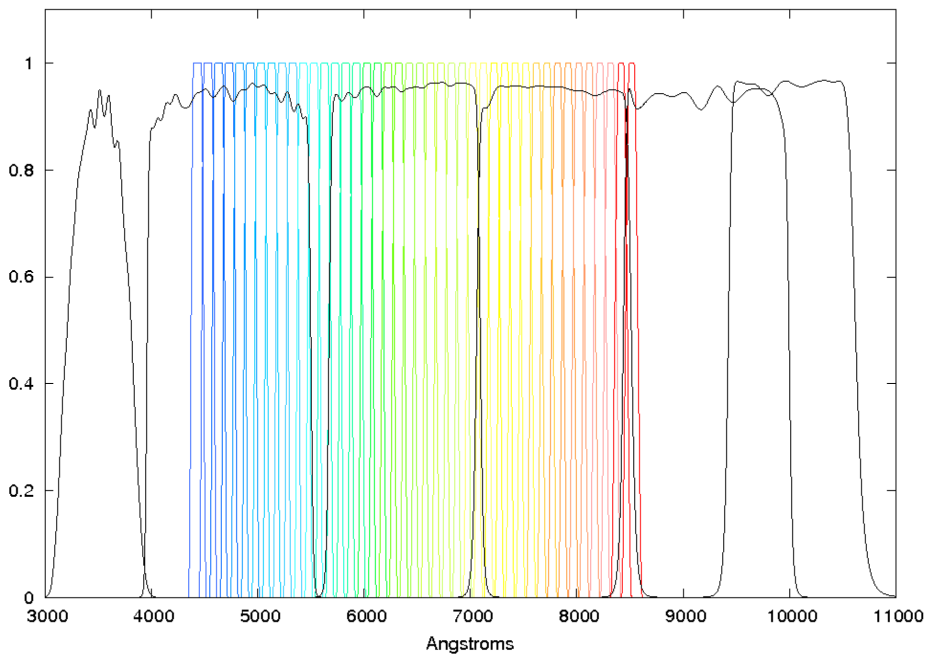

Figure 1. Wavelength ranges of the PAU filters, arbitrarily normalized.

space distortions and weak lensing, for which precise redshift measurements for large numbers of galaxies is imperative. For more details on the science enabled by this type of data set, see Gaztañaga et al. (2012). The survey plans to cover 100-200 square degrees using the PAUCam camera, which is currently under construction and will be mounted at the prime focus of the 4.2 meter William Herschel Telescope (WHT) at the Roque de los Muchachos Observatory on the island of La Palma[2].

PAUCam (Casas et al. 2011) includes 18 CCDs, each with 2048×4096 pixels of 15μ size. Each pixel subtends 0.26 arc seconds on the sky. The CCDs are laid out as shown in Figure 2, which also indicates that the field of view of PAUCam subtends one square degree. The field of view of the WHT is partially vignetted; the central 8 CCDs are mainly unvignetted, as 90% of the light is transmitted at the radius enclosing $40'$. The field is 50% vignetted at the radius enclosing 1°. Because of this vignetting, and because the telescope optics generate distortions in the outer regions of the field of view, we consider data from the central 8 CCDs as forming the main body of the PAU data set.

PAUCam will have multiple filter trays, arranged such that one filter covers each CCD. These filters will be housed in a jukebox-like system inside the cryostat, enabling the filters to be placed a few millimeters in front of the CCDs. Covering the 8 central CCDs with 40 narrow band + 6 broad band filters will require at least 6 filter trays.

PAU will form a powerful compromise between the redshift accuracy of a spectroscopic survey and the completeness and efficiency of an imaging survey. The photomet-

[2]http://www.ing.iac.es/Astronomy/telescopes/wht/

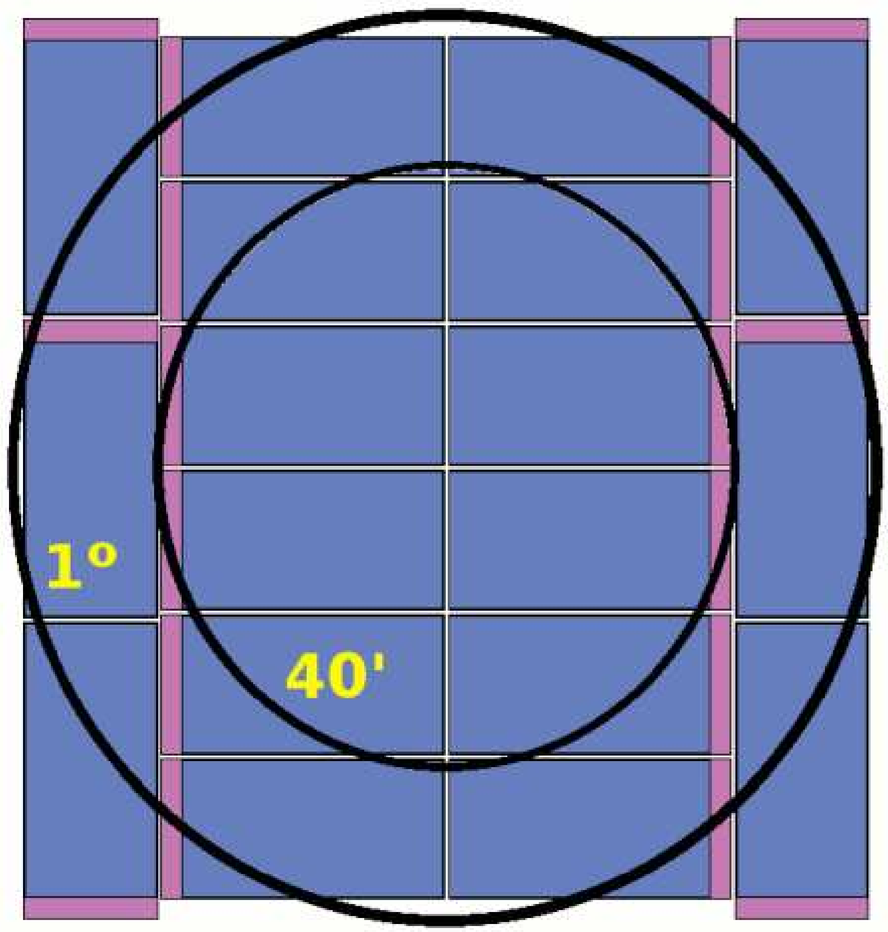

Figure 2. Layout of the 18 CCDs on PAUCam, with the size of the field of view superimposed.

ric redshift quality will enable precise analyses of galaxy clustering, without the need to impose strict selection criteria as is needed in typical spectroscopic projects. The success of the survey's science goals depends on the accuracy and precision of the redshift measurements. Because photometric redshifts are determined from galaxy colors, their quality is directly linked to the photometric calibration of the data. It is therefore imperative that the PAU Survey data are carefully calibrated, and that we thoroughly test the data reduction and calibration with realistic simulations.

2. PAU Data Management

The flow of data through the PAU data management pipelines is shown in Figure 3. The process begins with the arrival of data, or (as is the case until commissioning of the camera in 2013) with images produced by the Pixel Simulation pipeline. The raw data, real or simulated, are then given to the Nightly Processing pipeline, which produces level 1 products: clean images that are saved in storage, and object catalogs that are ingested into our database.

After the survey has observed for sufficient time to obtain multiple images over the same sky area in all of the PAU filters, the Multi-Epoch and Multi-Band (MEMBA) pipeline is run to coadd overlapping image data and measure photometry to the full depth of the available data. The level 2 products produced by the MEMBA pipeline consist of coadded images and final object catalogs, with one combined flux measurement per object per filter. These products are saved in storage and the database, as with the level 1 products.

After the Nightly or MEMBA pipelines are run, the results are checked using the Analysis pipeline, which ensures that the data quality is sufficient to be used for science.

The PAU data management pipelines are implemented at the Port d'Informació Cientfica (PIC)[3], which provides hardware and software support in the development, implementation, and execution of the pipelines, storage, and database.

2.1. Nightly Processing

The Nightly Processing pipeline is outlined in Figure 4. It follows standard data reduction techniques, and liberally uses publicly available Astromatic software[4] written by Emmanuel Bertin.

Raw image data from the telescope or Pixel Simulation pipeline are in the format of one FITS file per exposure, with one extension per amplifier ($\times$ 4 amplifiers on each of 18 CCDs = 72 extensions). The four amplifier images per CCD are corrected by their different gains, overscan corrected, and bias subtracted. The flat field input images are iteratively clipped to identify pixels with atypical response, which are added to a bad pixel mask. Cosmic rays are identified using the algorithm L.A. Cosmic (van Dokkum 2001). Saturated pixels are identified and added to the mask. The data are flat fielded to produce a clean image. The four amplifier regions per CCD are joined together to reconstruct one image per CCD. The resulting clean "science" image has 18 extensions,

[3] http://www.pic.es

[4] http://www.astromatic.net

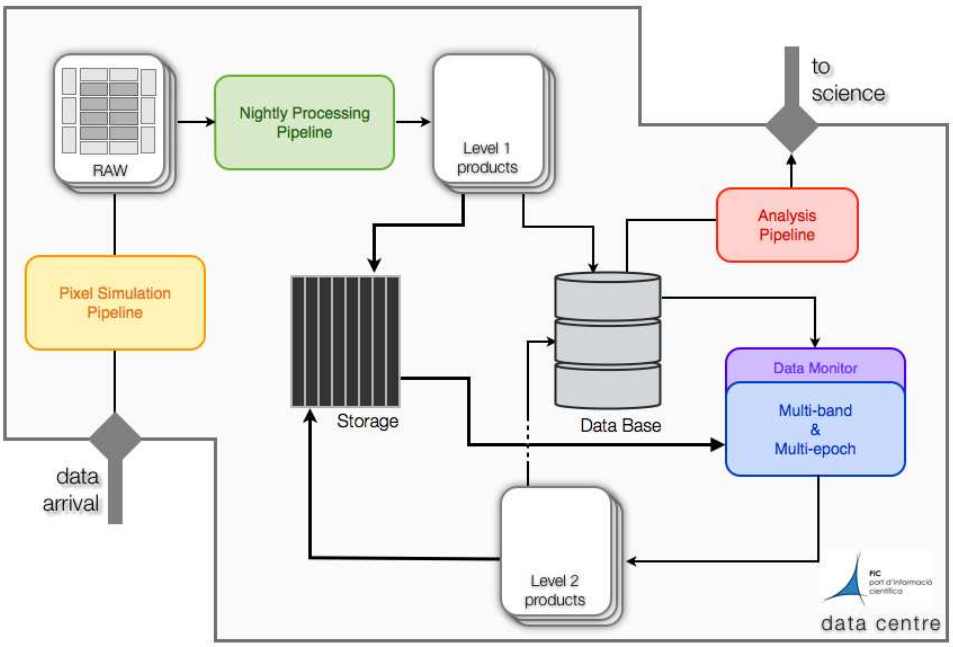

Figure 3. Flow chart of the PAU data management, showing different pipelines, storage, and data products.

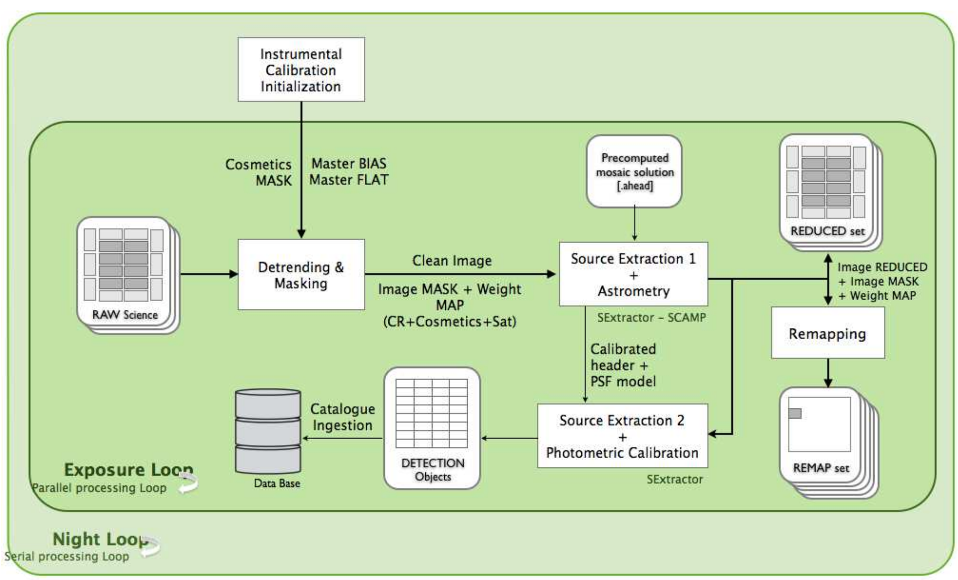

Figure 4. Flow chart of the PAU nightly pipeline.

and is accompanied by a mask and a weight map (which is given by the flat field, with bad pixels down-weighted by a large factor).

A high signal to noise detection catalog is extracted from the science image using SExtractor (Bertin & Arnouts 1996). This catalog is used to calculate an astrometric solution for the image using SCAMP (Bertin 2006), and also to characterize the image's point spread function (PSF) using PSFEx (Bertin 2011).

Using the astrometric solution and PSF model, SExtractor is run again with a lower signal to noise threshold to obtain the final detections catalog. The photometric calibration is derived using this catalog, as described in section 2.1.1. The final detections catalog and the photometric zero points are ingested into the database. The cleaned science image is saved in storage, along with its corresponding mask and weight map.

As a final step in the Nightly Processing pipeline, the cleaned images are remapped onto a fixed grid on the sky. The sky is divided into Fields, each one square degree, with remapped pixels the same size as the raw data pixels (0.26"). Each cleaned image is remapped into one or more "Tile" images, with one remapped Tile generated for each Field that the cleaned image overlaps. These remapped Tiles serve as inputs to the MEMBA pipeline.

2.1.1. Nightly Pipeline Calibration

Because the PAU filter system is unique, calibrating PAU data using existing surveys is not straightforward. The PAU Survey footprint will be covered by existing broadband surveys such as SDSS[5] or CFHTLens[6], which provide photometry in *ugriz* filters. The Nightly Processing pipeline calibration strategy is to convert these public broadband measurements into the PAU photometric system, to create a standard catalog of main sequence stars in PAU's 46 filters.

The "main sequence" refers to stars, fueled by hydrogen fusion at their core, that are in hydrostatic equilibrium. The colors of these stars are mainly determined by their mass, with small corrections due to other effects such as variations in metallicity. The colors of main sequence stars inhabit a compact region in color space, as they are close to those of black bodies. In particular, main sequence stars of a given spectral type have very similar spectra. Pickles (1998) has compiled a library of composite stellar spectra, combining data from different objects of the same spectral type in the wavelength range of 1150-25000Å. He notes that the RMS of the input spectra of the same spectral type is on the order of only 1-2%.

We use the spectral template library presented in Pickles (1998) to infer fluxes in the PAU filter system given measurements in broad band filters. In particular, in the case of overlap with SDSS, we convolve the stellar template spectra with the SDSS filter curves to generate SDSS-like colors of the template library. In the field of view of the given PAU exposure, we then query the SDSS database for well-measured stars. Each SDSS star is matched against the SDSS-like colors of the spectral library to find a best-fit stellar template spectrum. The best-fit stellar spectrum is then convolved with the 46 PAU filter curves to generate magnitudes of the stars in the PAU photometric system. This standard star catalog is then used to calculate a photometric zero point for each CCD on the focal plane.

[5]http://www.sdss.org

[6]http://www.cfhtlens.org

Advantages of this calibration method include the fact that a standard star catalog is generated including magnitudes in all filters, and therefore each CCD's data are compared independently to the catalog. This independence of the calibration across CCDs minimizes the correlation of photometric errors between different CCDs on the same exposure. Calibration errors that are correlated across the focal plane would lead to spatially correlated errors in galaxy color measurements, which would in turn cause spatially correlated photometric redshift errors that would degrade the galaxy clustering science analyses. Furthermore, since we can compare the data itself to standard measurements (rather than relying on standard star fields observed periodically during the night), we do not require the data to be photometric in order for the calibration to be accurate. The main source of systematic error in this method is inaccuracy in or poor matching with the stellar spectral templates. Given that the templates have been seen to be stable to 1-2%, and we expect on the order of $\gtrsim$50 stars matching with a standard survey on each chip, we do not expect template errors to be a dominant source of error in the PAU photometry. We plan to test this expectation in detail using the Pixel Simulations, which are described below.

Because the PAU filter system includes broad band *ugri*, which will be present in the survey that will serve as our calibration standard, we can calculate broad band zero points by simply comparing our data to the standard survey's, applying color terms which we expect to be minor and which we can quantify during the commissioning of the camera. A broad band zero point calculated in this manner can be extrapolated to find zero points for narrow band data from the same exposure by assuming the wavelength dependence of atmospheric extinction. The wavelength dependence has been calculated for the Roque de los Muchachos Observatory[7], although this calculation does not include the variable effects of water vapor that cause significant extinction in the redder bands. More detailed dependencies can be modeled, for example using MODTRAN[8], although at this time we do not plan to incorporate such models in the PAU photometric calibrations. Given a model of the wavelength dependence of extinction, we can calculate the narrow band filters' zero points from those of the broad bands without relying on the use of a stellar spectral template library. We expect the resulting zero points to have a larger error than those calculated using our primary calibration technique, since the stellar spectral templates are more precise than the atmospheric extinction curve calculated by the Observatory. Still, this technique serves as a consistency check of the calibration, helping to identify potential systematic errors.

2.2. Multi-Epoch Processing

The flow of the Multi-Epoch Multi-Band (MEMBA) pipeline is shown in Figure 5. Broad band data from the central eight CCDs are PSF homogenized and coadded, using SWarp (Bertin et al. 2002), to obtain a deep image for the purpose of detecting objects. This deep image is used in the dual-detection mode of SExtractor, where objects are detected on the deep "detection" image but their flux is measured on a different "measurement" image. For data taken with the central 8 CCDs, the cleaned remapped images in each band (created during the nightly processing) are coadded to create one measurement image in each of the 46 filters. The boundary CCDs are treated differently

[7] http://www.ing.iac.es/Astronomy/observing/manuals/ps/tech_notes/tn031.pdf

[8] http://www.modtran5.com/

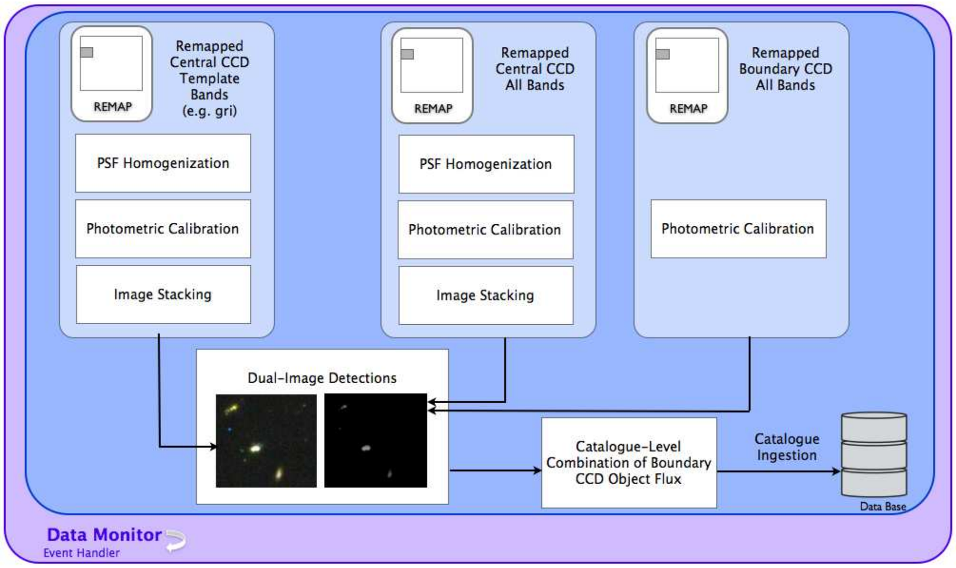

Figure 5. Flow chart of the PAU multi-epoch pipeline.

from the central 8; the outer regions of the focal plane suffer from vignetting and PSF distortions that would compromise the quality of a coadded image. We therefore use individual cleaned, remapped images from the boundary CCDs as measurement images, and combine the flux measurements with those from the coadded central 8 CCDs at the catalog level.

2.2.1. Multi-Epoch Calibration

The PAU Survey will observe its footprint at least twice in each filter. We can use the increased statistics given by multiple observations to improve upon the quality of the Nightly Processing pipeline calibration. In particular, we solve for a zero point for each image that minimizes the photometric offset between different observations over the same area, in the same filter. We follow the übercalibration procedure used in SDSS (Padmanabhan et al. 2008) and other recent surveys (e.g. Pan-STARRS; High et al. 2009). The statistics of the calibration are also greatly improved over that of the Nightly Processing pipeline if the Nightly calibration uses a standard survey that is shallower than PAU, such as SDSS.

As input to the übercalibration we use magnitudes measured on a single exposure and calibrated in the Nightly Processing pipeline, as stored in the Detection table in our database. These detections are matched by position, with the results of the matching stored as Global Objects in the database, providing easy references to measurements made of the same physical object in different exposures. These multiple measurements, in the same filter, are given as input to the übercalibration routine. The output is a set of zero points, one per image, which refine the zero points calculated in the Nightly Processing pipeline. These zero points are used in combination with those from the Nightly Processing pipeline to properly normalize and weight the images' Tiles during coaddition in the MEMBA pipeline.

The standard implementation of the übercalibration deals with only one filter's data at a time. Zero point shifts imposed by this step, then, are uncorrelated across the filters and may cause the colors of objects to become inaccurate. It is therefore important to check the object colors and introduce further shifts in the calibration zero points to correct them if necessary. Our primary technique is to do this by comparing the measured stellar colors to those of main sequence stars, in the spirit of High et al. (2009). We are also investigating the possibility of insisting upon proper colors of main sequence stars while performing the übercalibration, thereby imposing extra constraints on the zero points such that the resulting colors require no further shifts[9].

2.3. Analysis Pipeline

The Analysis Pipeline is a tool to ensure that the data quality meets the requirements necessary for science analyses. For the purposes of pipeline development using the image simulations, it compares the pipeline outputs against truth tables from which the simulations were generated. For use with either simulations or data, internal checks such as astrometric and photometric homogeneity are made, as well as measurements of detection completeness and purity against public surveys such as SDSS. The Analysis pipeline is currently used as a tool to evaluate improvements to the Nightly and MEMBA pipelines, with the ultimate goal of validating that the data management fully meets survey requirements.

2.4. Image Simulations

The Pixel Simulation pipeline is built around the software Skymaker (Bertin 2009), which takes input catalogs of stars and galaxies, and information about the survey camera and telescope, and outputs FITS images of a simulated sky. A schematic of this pipeline is shown in Figure 6.

As an input galaxy catalog we use results of the MICE simulation[10], which is a large cosmological N-body dark matter simulation that has been populated with galaxies using a halo occupation distribution model. This procedure results in a galaxy catalog with clustering properties that match those we expect to see in the real data set. Our use of the MICE catalogs as input allows us to run galaxy clustering science analyses on the output of the pipeline as run on the image simulations, so that we can test our entire set of data reduction and analysis software from beginning to end, ensuring that the data management is of sufficient quality to enable precision measurements of cosmology.

The star catalog given to Skymaker is generated from three sources: very bright stars (mag $\lesssim 15$) are taken from the USNO-B1.0 catalog (Monet et al. 2003), fainter stars (mag $\lesssim 21$) are taken from SDSS, and the faintest stars are generated using the Besançon model of the Milky Way (Robin et al. 2003). In this manner we include stars at all magnitude ranges expected in the PAU Survey, including bright stars in their correct positions that can be used for astrometry when comparing against reference catalogs such as SDSS or USNO.

---

[9]Thank you to Lynne Jones and Tim Axelrod for the suggestion!

[10]http://maia.ice.cat/mice

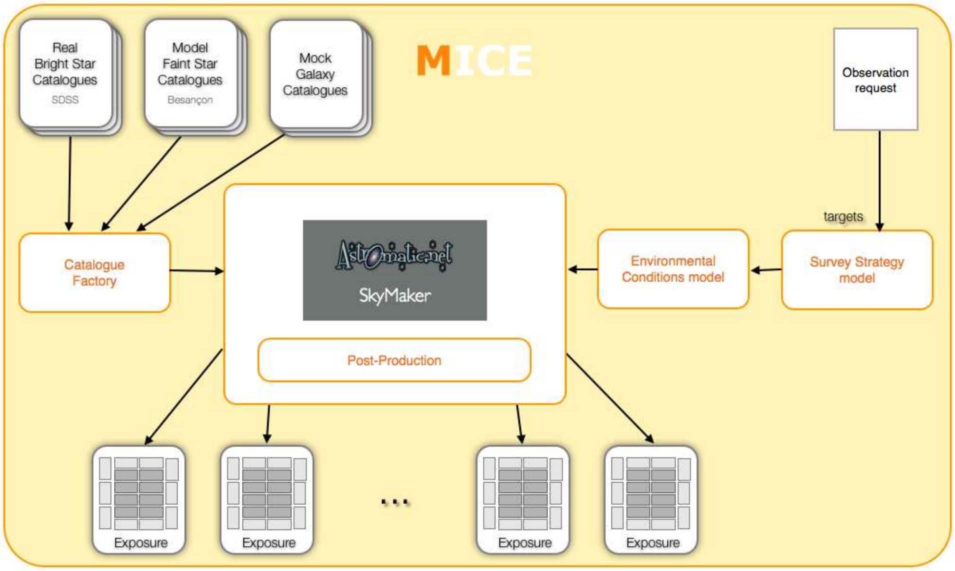

Figure 6. Flow chart of the PAU image simulation pipeline.

The output of the Pixel Simulation pipeline is one FITS file per exposure, with 72 extensions as will exist in the PAU data. The pipeline also produces bias and flat field images to be used as input to the Nightly Processing pipeline.

3. Status and Conclusion

To date, the pipelines described above are partially implemented. The fundamentals of the Nightly Processing pipeline are complete, able to run both on PCs and in parallel in the GRID framework at the PIC data center. The Pixel Simulation pipeline is also implemented, and is being used to test the Nightly pipeline. As elements of the Nightly pipeline are validated we are adding complexity to the image simulations (e.g. image defects, PSF distortions) such that we develop the two pipelines in tandem in a manner that is convenient for fully understanding and debugging the software. The MEMBA pipeline is in development. The Analysis pipeline is partially implemented, providing diagnostic plots regarding the quality of the Nightly pipeline output.

The photometric system of the PAU Survey is unique, combining 40 custom narrow band filters with 6 standard broad bands. This configuration allows for measurements of the galaxy redshifts that are an order of magnitude more precise than those measured using broad band data alone. These high-quality redshift measurements enable galaxy clustering analyses that will allow the PAU Survey to measure cosmological parameters with a precision competitive with other much larger surveys.

In order for the PAU Survey to measure high-quality redshifts, its photometric calibration must be accurate and robust. We have described the data reduction techniques used in the PAU data management pipelines, with emphasis on the photometric calibrations. We have also described the testing of these pipelines using custom image simulations based on N-body simulations, which allow for validation of the software all the way from the cleaning of the raw data, through both nightly and multi-epoch processing and calibrations, to the final outputs of the science analysis software. Using these techniques we can best control the quality of the PAU software, data products and science results.

Acknowledgments. Funding for this project was partially provided by the Spanish Ministerio de Ciencia e Innovacion (MICINN), project AYA2009-13936, Consolider-Ingenio CSD2007-00060, research project 2009-SGR-1398 from Generalitat de Catalunya, Marie Curie European Reintegration Grant PERG07-GA-2010-268290, and the CSIC JAE-Doc program.

References

Bertin, E. 2006, in Astronomical Data Analysis Software and Systems XV, edited by C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, vol. 351 of Astronomical Society of the Pacific Conference Series, 112

— 2009, Mem. Societa Astronomica Italiana, 80, 422

— 2011, in Astronomical Data Analysis Software and Systems XX, edited by I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, vol. 442 of Astronomical Society of the Pacific Conference Series, 435

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393

Bertin, E., Mellier, Y., Radovich, M., Missonnier, G., Didelon, P., & Morin, B. 2002, in Astronomical Data Analysis Software and Systems XI, edited by D. A. Bohlender, D. Durand, & T. H. Handley, vol. 281 of Astronomical Society of the Pacific Conference Series, 228

Casas, R., Ballester, O., Cardiel-Sas, L., Carretero, J., Castander, F. J., Castilla, J., Crocce, M., de Vicente, J., Delfino, M., Fernández, E., Fosalba, P., García-Bellido, J., Gaztañaga, E., Grañena, F., Jiménez, J., Madrid, F., Maiorino, M., Martí, P., Miquel, R., Neissner, C., Ponce, R., Sánchez, E., Serrano, S., Sevilla, I., Tonello, N., & Troyano, I. 2011, in Highlights of Spanish Astrophysics VI, edited by M. R. Zapatero Osorio, J. Gorgas, J. Maíz Apellániz, J. R. Pardo, & A. Gil de Paz, 674

Gaztañaga, E., Eriksen, M., Crocce, M., Castander, F. J., Fosalba, P., Marti, P., Miquel, R., & Cabré, A. 2012, MNRAS, 422, 2904. 1109.4852

High, F. W., Stubbs, C. W., Rest, A., Stalder, B., & Challis, P. 2009, AJ, 138, 110. 0903.5302

Monet, D. G., Levine, S. E., Canzian, B., Ables, H. D., Bird, A. R., Dahn, C. C., Guetter, H. H., Harris, H. C., Henden, A. A., Leggett, S. K., Levison, H. F., Luginbuhl, C. B., Martini, J., Monet, A. K. B., Munn, J. A., Pier, J. R., Rhodes, A. R., Riepe, B., Sell, S., Stone, R. C., Vrba, F. J., Walker, R. L., Westerhout, G., Brucato, R. J., Reid, I. N., Schoening, W., Hartley, M., Read, M. A., & Tritton, S. B. 2003, AJ, 125, 984. arXiv:astro-ph/0210694

Padmanabhan, N., Schlegel, D. J., Finkbeiner, D. P., Barentine, J. C., Blanton, M. R., Brewington, H. J., Gunn, J. E., Harvanek, M., Hogg, D. W., Ivezić, Ž., Johnston, D., Kent, S. M., Kleinman, S. J., Knapp, G. R., Krzesinski, J., Long, D., Neilsen, E. H., Jr., Nitta, A., Loomis, C., Lupton, R. H., Roweis, S., Snedden, S. A., Strauss, M. A., & Tucker, D. L. 2008, ApJ, 674, 1217. arXiv:astro-ph/0703454

Pickles, A. J. 1998, PASP, 110, 863

Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, A&A, 409, 523

van Dokkum, P. G. 2001, PASP, 113, 1420. arXiv:astro-ph/0108003